Science World Journal Vol. 19(No 1) 2024 https://dx.doi.org/10.4314/swj.v19i1.12
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Full Length Research Article

ESTIMATION AND COMPARISON OF WEIBULL-NORMAL DISTRIBUTION WITH SOME OTHER PROBABILITY MODELS USING BAYESIAN METHOD OF ESTIMATION

*Suleiman Aliyu and Muhammad Ardo Bamanga

Department of Mathematical Sciences, Kaduna State University, Kaduna

*Corresponding Author Email Address: suleimanaliyu992@gmail.com

ABSTRACT
In statistical applications the Normal distribution is adjudged to be the best. Recent studies Terna (2017) using classical method indicated that Weibull-Normal distribution outperformed the Normal distribution. In this study we used the non-classical Bayesian method of estimation to estimate and compare the Weibull-Normal distribution with some other distributions including Normal and Gamma-Normal distributions. This study derived explicit expressions for basic statistical properties such as moments, moment generating function, the characteristic function, reliability analysis and the distribution of order statistics. It looks at estimation of confidence intervals for the parameters of the Weibull-Normal distribution and estimated the parameters of the new distribution using a non-classical approach for the purpose of theoretical comparisons. The two other distributions whose parameters were also estimated by using Bayesian estimation are the normal distribution and gamma distribution as well as the combination Gamma-Normal distribution. Based on the analyses and interpretations of the results obtained it was found that the parameters and other general properties of Normal distribution gives a better fit than other distributions. R-software was used; the models were written as an R code in R program using the rjags library, the distribution parameters were obtained from a Gibbs sampling of a Bayesian Fit for data set I and data set II. The criteria used in R for comparisons were the negative log-likelihood, AIC (Akaike information criterion), CAIC (Consistent Akaike Information Criterion) and BIC (Bayesian information Criterion).

Keywords: Weibull-Normal Distribution, Bayesian Estimation, R-Programing software. Gibbs Sampling, AIC (Akaike information criterion), CAIC (Consistent Akaike Information Criterion) and BIC (Bayesian information Criterion)

INRODUCTION
Probability distribution and estimation theory are two important tools in Statistics for the study of random phenomena. In real life there is no certainty about what will happen in the future, but decisions still have to be made. Therefore, decision processes must be able to deal with the problems of uncertainty. Events that cannot be predicted precisely are often called random events. Many, if not most, of the input to, and processes that occur in, our systems are to some extent random. Hence, also the outputs or predicted impacts, and even people's reactions to those outputs or impacts? To ignore this randomness or uncertainty is to ignore reality. One of the commonly used tools for dealing with uncertainty in planning and management is probability. Probability is a branch of Mathematical Statistics that is used for quantitative modeling of random variables. The probability of an event represents the proportion of times under certain conditions that the outcome can be expected to occur. A probability density function is a mathematical description that approximately agrees with the frequencies or probabilities of possible events of a random variable.

Bayesian methods have become increasingly popular in analysis of geostatistical data in recent years. The Bayesian paradigm provides a coherent approach for specifying sophisticated hierarchical models for complex data, and recent computational advances have made model fitting in these situations feasible. The drawback to using Bayesian methods is that all joint distributions of processes and parameters have to be specified, often via a collection of conditional distributions. Distributional assumptions are not necessary in classical geostatistics, which require only the specifications, a powerful advantage of the Bayesian approach is that it allows a deeper scientific understanding of the nature of the process.

Many lifetime data used for statistical analysis follows a particular probability distribution and therefore knowledge of the appropriate distribution that any phenomenon follows greatly improves the sensitivity, reliability and efficiency of the statistical analysis associated with it.

Furthermore, it is true that several probability distributions exist for modeling lifetime data; however, some of these lifetime data do not follow any of the existing and well-known standard probability distributions (models) or at least are inappropriately described by them. This, therefore, creates room for developing new distributions and addresses the problem of estimation of parameters of some well-known distribution functions which could better describe some of these phenomena and therefore provide greater flexibility and wider acceptability in the modeling of lifetime data.

Gibbs sampling is a way of sampling from a probability distribution of two or more dimensions or multivariate distribution. It's a method of Markov Chain Monte Carlo which means that it is a type of dependent sampling algorithm. Because of that, if we need independent samples, some samples from the beginning are usually discarded because they may not accurately represent the desired distribution.

The main point of Gibbs sampling is that given a multivariate distribution, it's simpler to sample from a conditional distribution than from a joint distribution. For instance, instead of sampling directly from a joint distribution $P(x, y)$, Gibbs sampling propose sampling from two conditional distribution $P(x|y)$ and $P(y|x)$

For a joint distribution $P(x, y)$, we start with a random sample $(x^{(0)}, y^{(0)})$. Then we sample $x^{(1)}$ from the conditional distribution $P(x|y^{(0)})$. And $y^{(1)}$ from the conditional distribution $P(y|x^{(1)})$.

Weibull Distribution Function

The Weibull distribution has received much interest in reliability theory (Mahdi & Arjun, 2013). Murthy *et al.* (2004) stated that the use of Weibull distribution to describe real phenomena has a long history. According to Math page (2016), the Weibull distribution was originally proposed by the Swedish physicist Waloddi Weibull (1979) He used it for modeling the distribution of breaking strength

of materials. Since then, it has received applications in many areas. The Weibull distribution was named after engineer Waloddi Weibull (1887–1979) who popularized its use for reliability analysis, especially for metallurgical failure modes.
The Probability Distribution Function (PDF) of Weibull Distribution
The Weibull distribution with non-zero location/shift has three parameters, denoted as follows: $\gamma > 0$ is the shape parameter $\beta > 0$ is the scale parameter and α, the location or shift parameter which is also a lower bound. Its probability density function is given by (Denis, 2008):

$$f(x;\gamma,\beta,\alpha) = \gamma\beta^{-\gamma}(x-\alpha)^{\gamma-1}e^{-\left(\frac{x-\alpha}{\beta}\right)^{\gamma}} \quad (1)$$

When the location/shift parameter α is set to zero, the above equation (1) becomes two parameter Weibull distribution function with probability density function as:

$$f(x;\gamma,\beta,) = \gamma\beta^{-\gamma}x^{\gamma-1}e^{-\left(\frac{x}{\beta}\right)^{\gamma}} \quad (2)$$

WEIBULL-NORMAL DISTRIBUTION

For any continuous distribution with cumulative distribution function (CDF) $G(x,\zeta)$ and PDF, g(x,ζ) Bourguignon *et al*. (2014) proposed the Weibull generalized (denoted "Weibull-G) family of distributions that provides greater flexibility in modeling of real data sets. The cumulative distribution function (CDF) of the Weibull generalized family of distributions according to Bourguignon et al. (2014) is defined as

$$F_{WG}(x;\alpha,\beta,\zeta) = \int_{-\infty}^{\frac{G(x;\zeta)}{G'(x;\zeta)}} \alpha\beta t^{\beta-1}e^{-\alpha t^{\beta}}dt \quad (3)$$

Where $G(x,\zeta)$ is the cumulative distribution function (CDF) of any continuous distribution, which depends on the parameter vectorζ,$G`(x,\zeta) = 1 - G(x,\zeta) and \alpha > 0\ and \beta > 0$ are the scale and shape parameters respectively
The interpretation of the above Weibull generator is given as follows:
Let Y be a lifetime random variable having a certain continuous$G(x,\zeta)$ distribution. The odds ratio that an individual (or component) following the lifetime Y will die (failure) at time x is $G(x,\zeta)/G`(x,\zeta)$.If the variability of these odds of death is represented by the random variable X and assumed to follows a Weibull distribution with probability density function (PDF)$f(x) = \alpha\beta t^{\beta-1}e^{-\alpha t^{\beta}}$ where α is the scale parameter and β is the shape parameter. Hence, we can represent the above interpretation as

$$P(Y \le x) = P\left(X \le \frac{G(x;\zeta)}{G'(x;\zeta)}\right) = F(x;\alpha,\beta,\zeta) = \int_{-\infty}^{\frac{G(x;\zeta)}{G'(x;\zeta)}} \alpha\beta t^{\beta-1}e^{-\alpha t^{\beta}}dt \quad (4)$$

Using integration by substitution in equation (4), we perform the following operations:
Let $u = e^{-\alpha t^{\beta}} = e^{-v}$ where$v = \alpha t^{\beta}$then

$$F_{WG}(x;\alpha,\beta,\zeta) = \int_{-\infty}^{\frac{G(x;\zeta)}{G'(x;\zeta)}} \alpha\beta t^{\beta-1}u dt \quad (5)$$

$$\frac{du}{dt} = \frac{du}{dv} \times \frac{dv}{dt}(thechainrule)$$

Where $\frac{du}{dv} = -e^{-v}$ and $\frac{dv}{dt} = \alpha\beta t^{\beta-1}\frac{du}{dt} = \frac{du}{dv} \times \frac{dv}{dt} = -\alpha\beta t^{\beta-1}e^{-\alpha t^{\beta}}$

$$dt = -\frac{du}{\alpha\beta t^{\beta-1}e^{-\alpha t^{\beta}}} = -\frac{du}{\alpha\beta t^{\beta-1}u} \quad (6)$$

Therefore, substituting for dt in equation (5) and simplifying, we obtain

$$F_{WG}(x;\alpha,\beta,\zeta) = -\int_{-\infty}^{\frac{G(x;\zeta)}{G'(x;\zeta)}} du \quad (7)$$

$$F_{WG}(x;\alpha,\beta,\zeta) = -[u]_{-\infty}^{\frac{G(x;\zeta)}{G'(x;\zeta)}} = -\left[e^{-\alpha t^{\beta}}\right]_{-\infty}^{\frac{G(x;\zeta)}{G'(x;\zeta)}} \quad (8)$$

The odd ratio that an individual following the lifetime Y will die at time x is $G(x,\zeta)/G'(x,\zeta)$where $G'(x,\zeta) = 1 - G(x,\zeta)$. If x assumes a Weibull distribution as in (5), then as $x \to -\infty$, $t = G(x,\zeta)/G'(x,\zeta)$will tend to $t = G(-\infty;\zeta)/G'(-\infty,\zeta) = 0$
Evaluating the integrand in equation (8) yields

$$F_{WG}(x;\alpha,\beta,\zeta) = 1 - \exp\left\{-\alpha\left[\frac{G(x;\zeta)}{G'(x;\zeta)}\right]^{\beta}\right\}, -\infty < x < \infty, \alpha, \beta > 0 \quad (9)$$

Therefore, equation (9) is the cumulative distribution function (CDF) of the Weibull-G family of distributions proposed in Bouguignon *et al*. (2014) and the corresponding PDF of the Weibull-G family can be obtained from equation (9) by taking the derivative of the cumulative distribution function (CDF) with respect to x as follows:

$$f(x) = F_{WG}(x;\alpha,\beta,\zeta) = \frac{d(F_{WG}(x;\alpha,\beta,\zeta))}{dx} \quad (10)$$

Let $y = \exp\left\{-\alpha\left[\frac{G(x;\zeta)}{G'(x;\zeta)}\right]^{\beta}\right\} = e^{-u}$

$$u = at^{\beta} and t = \frac{G(x;\zeta)}{G'(x;\zeta)} such that y = e^{-at^{\beta}} = e^{-u}$$

Where $F_{WG}(x;\alpha,\beta,\zeta) = 1 - \exp\left\{-\alpha\left[\frac{G(x;\zeta)}{G'(x;\zeta)}\right]^{\beta}\right\} = 1 - y$

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dt} \times \frac{dt}{dx}(funtionofafunctionruleorthechainrule)$$

Then,

$$\frac{dy}{du} = -e^{-u}, \frac{du}{dt} = \alpha\beta t^{\beta-1} and \frac{dt}{dx} = \frac{g(x,\zeta)}{[G'(x;\zeta)]},$$

Note that $t = \frac{g(x,\zeta)}{G'(x;\zeta)} = \frac{u}{v}$and by quotient rule,

$$\frac{dt}{dx} = \frac{V.\frac{du}{dx} - U.\frac{dy}{dx}}{V^2} = \frac{[1 - G(x;\zeta)]g(x;\zeta) - G(x;\zeta)[-g(x;\zeta)]}{[G'(x;\zeta)]^2} = \frac{g(x;\zeta)}{[G'(x;\zeta)]^2}$$

https://dx.doi.org/10.4314/swj.v19i1.12

$$f(x) = F_{WG}(x;\alpha,\beta,\zeta) = \frac{d(F(x;\alpha,\beta,\zeta))}{dx} = -\frac{dy}{dx}$$

Where

$$-\frac{dy}{dx} = -\left\{-\alpha\beta t^{\beta-1}\frac{g(x;\zeta)}{[G'(x;\zeta)]^2}e^{-u}\right\}$$

$$-\frac{dy}{dx} = \alpha\beta\frac{g(x;\zeta)}{[G'(x;\zeta)]^2}\left(\frac{G(x;\zeta)}{G'(x;\zeta)}\right)^{\beta-1}\exp\left\{-\alpha\left[\frac{G(x;\zeta)}{G'(x;\zeta)}\right]^{\beta}\right\} \quad (11)$$

Hence,

$$f(x) = F_{WG}(x;\alpha,\beta,\zeta) = \alpha\beta g(x;\zeta)\frac{(G(x,\zeta))^{\beta-1}}{(G'(x;\zeta))^{\beta+1}}\exp\left\{-\alpha\left[\frac{G(x;\zeta)}{G'(x;\zeta)}\right]^{\beta}\right\} \quad (12)$$

Finally, equation (12) is the probability density function (PDF) of the Weibull-G family of distributions where $g(x;\zeta)$ and $G(x;\zeta)$ are the PDF probability density function and the cumulative distribution function (CDF) of any baseline continuous distribution respectively which depends on the parameter vector $\zeta = (\mu,\sigma)$ where $\alpha > 0$ $and \beta > 0$ are the scale and shape parameters respectively

Equation (12) is the PDF of any Weibull-G family of distributions and is most tractable when both cumulative distribution function (CDF) and PDF have simple analytic expressions. The major benefit of equation (12) is to offer more flexibility to extremes of the PDFs and therefore it becomes suitable for analyzing data with high degree of asymmetry.

Henceforth, let G be a continuous baseline distribution, for each G distribution, we define the Weibull-G distribution with two extra parameters α and β by following equation (12). A random variable X with PDF equation (12) is denoted by X~Weibull-G(α,β,ζ). The additional parameters induced by the Weibull generator above are sought as a vehicle to furnish a more flexible distribution. This class of distribution becomes a special case of the exponential generator if$\beta = 1$.

The CDF and PDF of the Normal distribution with location parameter $-\infty < \mu < \infty$ and scale parameter $\alpha > 0$ is given by

$$G(x;\mu,\sigma) = \Phi\left(\frac{x-u}{\alpha}\right) = \int_{-\infty}^{\frac{x-u}{\sigma}}\frac{1}{\sqrt{2\pi}}exp\left\{-\frac{(v-\mu)^2}{2\sigma^2}\right\}dv \quad (13)$$

And

$$g(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{\frac{(x-\mu)^2}{2\sigma^2}\right\} = \frac{1}{\sigma}\Phi\left(\frac{x-\mu}{\sigma}\right) \quad (14)$$

Where $\Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$.

Taking the CDF equation (13) and the PDF equation (14) of the Normal distribution with location parameter, $-\infty < \mu < \infty$ and dispersion parameter $\sigma > 0$.

The Cumulative Distribution Function (CDF) and Probability Distribution Function (PDF) of the Weibull-Normal Distribution (WND) are obtained from equation (10) and equation (11) as

$$F(x) = F(x;\alpha,\beta,\mu,\sigma) = 1 - \exp\left\{-\alpha\left[\frac{\phi\left(\frac{x-\mu}{\sigma}\right)}{1-\phi\left(\frac{x-\mu}{\sigma}\right)}\right]^{\beta}\right\}, -\infty < x < \infty, \alpha,\beta > 0$$

And $F(x) = F(x;\alpha,\beta,\mu,\sigma) =$

$$= \frac{\alpha\beta}{\sigma}\Phi\left(\frac{x-\mu}{\sigma}\right)\frac{\left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{\beta-1}}{\left[1-\phi\left(\frac{x-\mu}{\sigma}\right)\right]^{\beta+1}}\exp\left\{-\alpha\left[\frac{\phi\left(\frac{x-\mu}{\sigma}\right)}{1-\phi\left(\frac{x-\mu}{\sigma}\right)}\right]^{\beta}\right\}$$

Where$-\infty < X < \infty, \sigma > 0$ is the dispersion parameter, $-\infty < \mu < \infty$isthe location parameter and α>0 and β>0 are the scale and shape parameters respectively

Where,

$$\eta_r = \sum_{j,k=0}^{\infty} w_{j,k}S_{r(\beta(k+1)+j)}$$

By differentiating equation (15) and changing indices, we can obtain the PDF of the Weibull Normal distribution as

$$f(x) = \frac{dF(x)}{dx} = \sum_{r=0}^{\infty}\eta_r\, \mathrm{h}(\mathrm{x};\mathrm{r},\mu,\sigma)$$

$$f(x) = \sum_{r=0}^{\infty}\eta_r\,\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\frac{x-\mu}{\sigma}\right)^{r-1}$$

Where

$$\eta_r = \sum_{j,k=0}^{\infty} w_{j,k}S_{r(\beta(k+1)+j)} and \sum_{j,k=0}^{\infty}\eta_r = 1$$

Equation (16) is the PDF of the Weibull-Normal distribution (WND) defined as a linear combination of Exponentiated Normal Probability Density Function PDF.

ESTIMATION OF PARAMETERS OF THE WEIBULL-NORMAL DISTRIBUTION

Let $X_1, X_2, \dots, X_n$, be a random sample taken from Weibull-Normal distribution with unknown vector parameter$\theta = (\alpha,\beta,\mu,\sigma)^T$. The total log-likelihood function for θ is obtained from $f(x)$ as follows:

$$f(x) = \frac{\alpha\beta}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\frac{\left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{\beta-1}}{\left[1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{\beta+1}}\exp\left\{-\alpha\left[\frac{\Phi\left(\frac{x-\mu}{\sigma}\right)}{1-\Phi\left(\frac{x-\mu}{\sigma}\right)}\right]^{\beta}\right\} \quad (17)$$

The likelihood function is given by;

$$L(X_1,X_2,\dots,X_n/\alpha,\beta,\mu,\sigma) = \left(\frac{\alpha\beta}{\sigma}\right)^n\sum_{i=1}^{n}\phi\left(\frac{x_i-\mu}{\sigma}\right)\frac{\sum_{i=1}^{n}\left[\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{\beta-1}}{\sum_{i=1}^{n}\left[1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{\beta+1}}\exp\left\{-\alpha\left[\frac{\Phi\left(\frac{x_i-\mu}{\sigma}\right)}{1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)}\right]^{\beta}\right\} \quad (18)$$

Let the log-likelihood function, $l = logL(X_1,X_2,\dots,X_n/\alpha,\beta,\mu,\sigma)$.

Therefore

$$l = nlog\alpha + nlog\beta - nlog\sigma - \frac{N}{2}\log(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}(x_i-\mu)^2 + (\beta-1)\sum_{i=1}^{N}\log\left[\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right] - (\beta + 1)\sum_{i=1}^{N}\log\left[1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right] - \alpha\sum_{i=1}^{N}\left[\frac{\Phi\left(\frac{x_i-\mu}{\sigma}\right)}{1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)}\right]^{\beta}$$

Using equation (17) and equation (18) the joint uniform prior for the Weibull-Normal Distribution:

Science World Journal Vol. 19(No 1) 2024
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

https://dx.doi.org/10.4314/swj.v19i1.12

$$g(\mu) = p_j(\alpha,\beta) \propto \frac{1}{\alpha\beta}\left(\frac{1}{\sqrt{2\pi}}\right)e^{-\frac{\mu^2}{2}} \quad (19)$$

Using equation (18) and equation (19) the yield the posterior distribution of Weibull-Normal Distribution

$$p(\alpha,\beta|X) = \frac{1}{k}\frac{\alpha^{n-1}}{\beta^{n\alpha+1}}\prod_{i=1}^{n} x_i^{\alpha} e^{\left(-\frac{\sum_{i=1}^{n}\alpha x_i^{\alpha}}{\beta^{\alpha}}\right)} \frac{(n+1)^{\frac{1}{2}}}{\sqrt{2\pi}} e^{\left(-\frac{(n+1)}{2}\left(\mu-\frac{nX}{n+1}\right)\right)^2} \quad (20)$$

Maximization of equation (20) can be performed by using well established routines like the nlm (non-linear minimization) routine or optimize in the R Statistical package. Setting these equations to zero, and solving them simultaneously yields Bayesian estimation for the Weibull-Normal distribution. These equations cannot be solved analytically and therefore Statistical software can be used to solve them numerically by means of iterative techniques like the Newton-Raphson algorithm.

APPLICATIONS, RESULTS AND DISCUSSION

This section presents the results from the analysis for the Bayesian Method of Estimation for the Weibull-Normal distribution and discussion of the results. Weibull-Normal Distribution functions used for the research work are compared based on it fitted outcome using R Programming software.
Bayesian estimation technique thrives on the fundamental theorem of Bayes which basically is expressed as follows:

Posterior Distribution = Likelihood Distribution x Prior Distribution
For Weibull-Normal Distribution, the prior distribution which is the distribution of the parameters is a Normal distribution while the likelihood distribution is the Weibull distribution making the posterior distribution a Weibull-Normal Distribution. In other words, Equation 4.1 can be written more specifically for this study as follows:

Weibull Normal Distribution Posteror **= Weibull Distribution** Likellyhood **x Normal Distribution** Prior

Given that the probability density function (PDF) and cumulative distribution function (CDF) of the Normal distribution does not have a closed form analytical solution, it follows that Weibull Normal Distribution also lacks a closed form solution as well. This made it very difficult to carry out the required Bayesian estimation of its parameters analytically making the numerical approach the best option for Bayesian estimation of distribution parameters. One primary process of Bayesian method of estimation is called Gibbs Sampling which was used for this study. Since the computation was done numerically, R-programming software was used with aid of a library routine called rjags (just another Gibbs sampling). The next section describes the data set to be used to verify the validity of the Weibull-Normal Distribution in R programming.

Preliminary Analysis and Description

This study adopted two data sets to be used for Bayesian estimation of the parameters of Weibull Normal Distribution. The two data sets are described as follows:
Data set I: This data is on the strength of 1.5cm glass fibers. The data was originally obtained by workers at the UK national physical laboratory and it has been used by Smith and Naylor (1987), Barreto-Souza *et al.* (2011), Bourguignon et al. (2014), Oguntundeet*al*. (2015), Afify and Aryal (2016) and Terna (2017). This data can be found in Table 4.1 as follows:

Table 1: Data set I

0.74	0.77	0.81	0.84	0.93	0.104	1.11	1.13	1.24	1.25	1.271.28	1.29	
1.36	1.39	1.42	1.48	1.51	1.52	1.53	1.54	1.55	1.55	1.58	1.61	1.62
1.63	1.64	1.66	1.66	1.66	1.67	1.68	1.68	1.69	1.70	1.70	1.73	1.76 1.77
1.78	1.81	1.82	1.84	1.84	1.89	2.00	2.01	2.24				

Data set II: The second data set represents 66 observations of the breaking stress of carbon fibres of 50mm length (in GPa) given by Nicholas and Padgett (2006). This data set has been used by Cordeiro and Lemonte (2011), Al-Aqtashet *al*. (2014), Afify et *al*. (2014), Oguntunde *et al*. (2015), Afify et *al*. (2016), Terna (2017). This data can be found in Table 4.2. as follows.

Table 2: Data set II

, 1.87, 2.53, 2.85, 3.15, 3.56, 0.85, 1.89, 2.55, 2.87, 3.19, 3.60, 1.08, 2.03, 2.55, 2.88,
, 3.65, 1.25, 2.03, 2.56, 2.93, 3.22, 3.68, 1.47, 2.05, 2.59, 2.95, 3.27, 3.70, 1.57, 2.12,
, 2.96, 3.28, 3.75, 1.61, 2.35, 2.73, 2.97, 3.31, 4.20, 1.61, 2.41, 2.74, 3.09, 3.31, 4.38,
, 2.43, 2.79, 3.11, 3.33, 4.42, 1.80, 2.48, 2.81, 3.11, 3.39, 4.70, 1.84, 2.50, 2.82, 3.15,
, 4.90

The summary descriptive Statistics of the dataset is given in Table 3 below:

Table 3: Summary of the two data sets

ameters	a set I	a set II
imum	4	
Q_1	5	8
lian	0	5
Q_3	5	8
n	7	0
imum	0	0
ance	5	5
wness	79	29
osis	4	3

For clearer visualization of the data sets, histogram plots of the two data sets were made and are shown below in Figures 1 and 2:

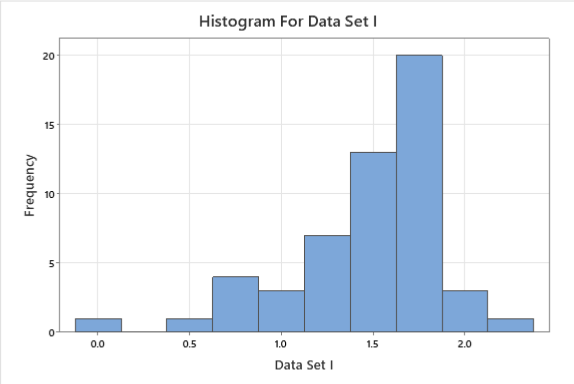

Figure 1: Histogram for Data Set I

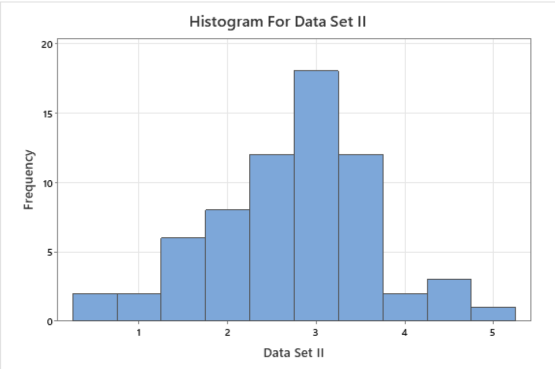

Figure 2: Histogram for Data Set II

Both the summary and the histograms for the two data sets show that the first data set is skewed and therefore, suitable for skewed or asymmetric distributions which the Weibull-Normal distribution is not an exception, while the second data set is approximately Normal which might not be analyzed adequately using asymmetric distributions like the Weibull-Normal distribution. There are two other distributions whose parameters were also estimated by using Bayesian parameter estimation such as Normal distribution and gamma distribution. These distributions were so chosen because of the data set available. Other distributions like Beta distribution and Kumaraswamy Distribution could not be used because of the type of data set that is most of the data set values are greater than 1, while distributions like Binomial, Bernoulli, Hyper-geometric, Poisson etc. it's not suitable because data set is continuous while the distributions are for discrete data set.

To compare these distributions, we will consider some criteria: the AIC (Akaike Information Criterion), CAIC (Consistent Akaike Information Criterion) and BIC (Bayesian Information Criterion). These Statistics are given as:

$$AIC = -2ll + 2k \quad (4.3)$$

$$BIC = -2ll + k \log n \quad (4.4)$$

$$CAIC = -2ll + \frac{2kn}{n-k-1} \quad (4.5)$$

Where *ll* denotes the log-likelihood function evaluated at Bayesian method of estimation, k is the number of the model parameters and n is the sample size. The model with the lowest values for these Statistics was chosen as the best model to fit the data.

Bayesian Parameter Estimation

Bayesian Parameter Estimation for Data Set I

After writing the given models as an R code in R program using the rjags library, the distribution parameters were obtained from a Gibbs sampling of a Bayesian Fit for data set I and presented as shown below in Table 4.4. From the table below Normal distribution was the most appropriate fit for the distribution followed by Weibull-Normal while Gamma distribution came last.

Comparison of the negative log-likelihood, AIC, CAIC, and BIC values for Normal distribution and Weibull-Normal distribution shows that their performance in relation to predicting the outcomes of the data set is nearly the same. Only Gamma-Normal distribution showed a large deviation in terms of performance criteria making it less suitable.

Science World Journal Vol. 19(No 1) 2024 https://dx.doi.org/10.4314/swj.v19i1.12
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Table 4: Performance of the distribution using the AIC, CAIC and BIC values of the models based on Data Set I (strength of 1.5cm glass fibers).

tributions	**lihood tribution**	**r**	**ameter mates**	**(minus log-lihood)**		**C**		**ks of del's formance**
bull-Normal	bull	mal	$\hat{\alpha} = 4.349$ $\hat{\beta} = 0.135$ $\hat{\mu} = 1.470$ $\hat{\sigma} = 0.055$	13	26	66	67	
mal	mal	mal	$\hat{\hat{\alpha}} = 1.481$ $\hat{\hat{\beta}} = 5.755$	69	38	78	79	
nma-Normal	nma	mal	$\hat{\alpha} = 5.812$ $\hat{\hat{\beta}} = 4.071$ $\hat{\mu} = 6.950$	40	79	19	20	

Science World Journal Vol. 19(No 1) 2024
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University
https://dx.doi.org/10.4314/swj.v19i1.12

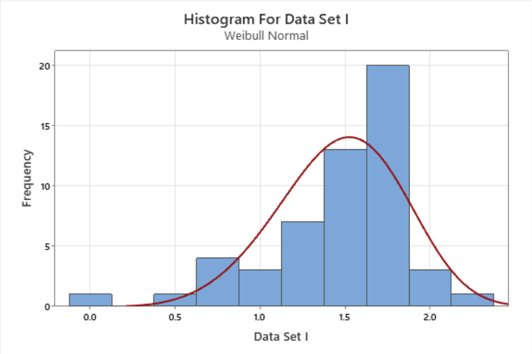

Figure 3: Fitting the Weibull-Normal Distribution to Histogram Data

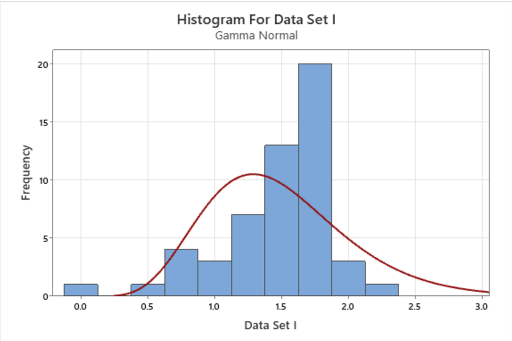

Figure 4: Fitting the Gamma-Normal Distribution to Histogram Data Set II

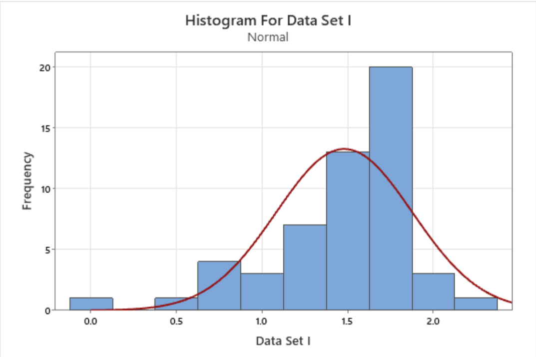

Figure 5: Fitting the Normal Distribution to Histogram Data Set I

A figure 4.3, 4.4 and 4.5 shows the level of agreement that exists between the data set and proposed distribution. The results still showed that Normal distribution fits best followed by Weibull Normal and lastly Gamma-Normal distribution.

4.3.2 Bayesian Parameter Estimation for Data Set II

The distribution parameters were obtained from a Gibbs sampling of a Bayesian Fit for data set II and presented as shown below in Table 4.5. From the table below Normal distribution was the most appropriate fit for the distribution followed by Weibull-Normal while Gamma distribution came last. Comparison of the negative log-likelihood, AIC, CAIC, and BIC values for Normal distribution and Weibull-Normal distribution shows that their performance in relation to predicting the outcomes of the data set is nearly the same. Only Gamma Normal distribution showed a large deviation in terms of performance criteria making it less suitable.

Figures 4.6, 4.7 and 4.8 show the level of agreement that exists between the data set and proposed distributions. The results still showed that Normal distribution fits best followed by Weibull-Normal distribution and lastly Gamma-Normal distribution.

https://dx.doi.org/10.4314/swj.v19i1.12

Table 4.5: Performance of the distribution using the AIC, CAIC and BIC values of the models based on data set II (breaking stress of carbon fibres of 50mm length in GPa).

tributions	lihood tribution	r	ameter mates	(minus log-lihood)		C		ks of model's formance
bull-mal	bull	mal	$\hat{\alpha} = 3.168$ $\hat{\hat{\beta}} = 0.033$ $\hat{\sigma} = 2.729$ $\hat{\sigma} = 0.114$	72	.944	.134	.323	
mal	mal	mal	$\hat{\hat{\sigma}} = 2.731$ $\hat{\hat{\beta}} = 1.199$	34	.069	.259	.449	
nma-mal	nma	mal	$\hat{\alpha} = 6.201$ $\hat{\hat{\beta}} = 2.270$ $\hat{\sigma} = 1.198$	51	.902	.093	.282	

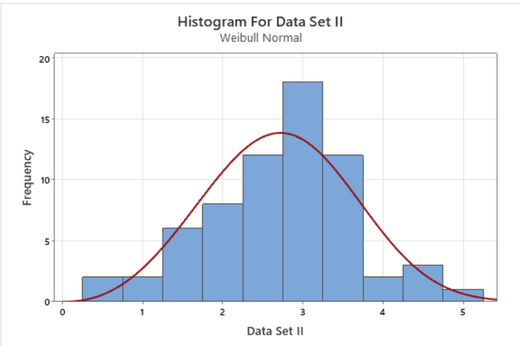

Figure 4.6: Fitting the Weibull-Normal Distribution to Histogram Data

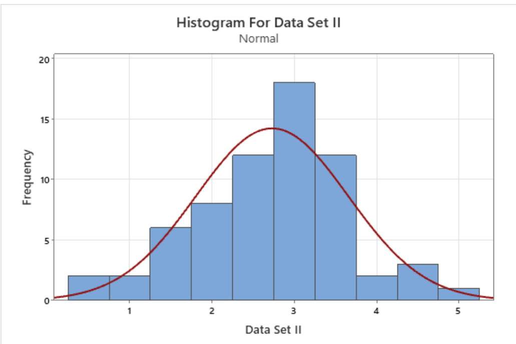

Figure 4.7: Fitting the Normal Distribution to Histogram Data Set II

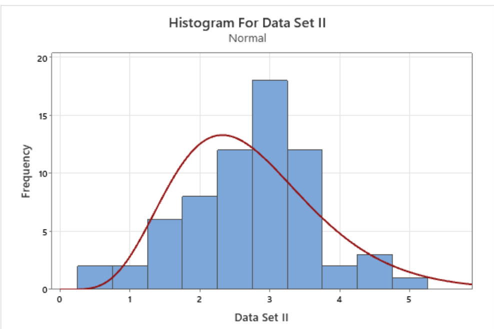

Figure 4.8: Fitting the Gamma-Normal Distribution to Histogram Data Set II Histogram Data Set II

DISCUSSION

Based on the analysis and interpretations of the results obtained it was found that the parameters of Weibull-Normal distributions obtained the general Statistics of the analyzed data that comprises of the Maximum, Minimum, Mean, Median, 1st Quartile, 3rd Quartile Variance, Skewness and Kurtosis respectively. After writing the given models as an R code in R program using the rjags library, the distribution parameters were obtained from a Gibbs sampling of a Bayesian Fit for data set I and data set II. From the result Normal distribution is the most appropriate fit for the distribution followed by Weibull-Normal while Gamma distribution came last. Comparison of the negative log-likelihood, AIC, CAIC, and BIC values for Normal distribution and Weibull-Normal distribution. It shows that their performance in relation to predicting the outcomes of the data set is nearly the same. Only Gamma-Normal distribution showed a large deviation in terms of performance criteria making it less suitable. The graphical representation of the fitted Weibull-Normal distribution shows the histogram of observed Normal distribution was the most appropriate fit for the distribution followed by Weibull-Normal distribution while Gamma Distribution came last.

The Weibull-Normal distribution density function was expressed as a linear combination of exponentiated Normal density function which allowed us to derive some of its mathematical properties like its ordinary and incomplete moments, moment generating function, characteristics function, reliability analysis and order Statistics. The estimation of parameters has been done using the non-classical approach (Bayesian Method of Estimation). The usefulness of the Weibull-Normal distribution has been illustrated by some applications to two real data sets. The results showed that the Normal distribution performs very well albeit slightly lesser than the Normal distribution for the both data sets. The difference in performance of the Normal distribution and Weibull-Normal distribution is quite negligible. However, the results confirmed that this distribution is more flexible and appropriate for modeling negatively skewed data sets.

Conclusion

In this research, a new four-parameter model called the Weibull-Normal distribution (WND) has been modified. Some mathematical and Statistical properties of the proposed distribution have been studied appropriately. An explicit expression has been derived for its moments, moment generating function, characteristics function, survival function, hazard function, and orderStatistics. Some plots of the distribution revealed that it is a skewed distribution and has only one mode. The model parameters have been estimated using the method of Bayesian parameter estimation. The study has proposed reliability measures widely used in engineering using the Weibull-Normal distribution.

The implications of the plots for the survival and hazard functions indicate that the Weibull-Normal distribution would be appropriate in modeling time or age-dependent events, where survival and failure rate decrease with time or age.

REFERENCES

Afify, A. Z. and Aryal, G. (2016). The Kumaraswamy Exponentiated Frechet Distribution. Journal of Data Science, 6, 1-19.

Afify, A.Z., Nofal, W. M. and Butt, N. S. (2014). Transmuted Complementary Weibull Geometric distribution. Pakistan Journal of Statistics and Operation Research, 10: 435-454.

Al-Aqtash, R., Lee, C. and Famoye, F. (2014). Gumbel-Weibull distribution: Properties and applications. Journal of Modern Applied Statistical Methods, 13: 201-225.

Barreto-Souza, W. M., Cordeiro, G. M. and Simas, A. B. (2011). Some results for beta Frechet distribution. Communication in Statistics: Theory and Methods. 40:798-811

Bourguignon, M., Silva, R. B. and Cordeiro, G. M. (2014). The Weibull-G Family of Probability Distributions. Journal of Data Science, 12, 53-68.

Cordeiro G. and Lemonte (2011). The Mcdonald Normal(MCN) Distribution. Pakistan Journal of Statistics and Operation Research, 8: 301-329.

Denis, C. (2008). Fitting the Three-Parameter Weibull Distribution: Review and Evaluation of

Mahdi M. and Arjun, (2013).Probability Distribution and Bayesian estimation of two unknown parameters of the Weibull distribution.International Journal of Scientific Research in Mathematical and Statistical Sciences

Murthy I, Srivastava, D. K., and Freimaer, M. (2004). Generalized Weibull Family: A Structural Analysis. Communication Statistics: Theory and Methods, 23(4): 1149-1171.

Nichols, M. D. and Padgett,W. J. (2006). A bootstrap control chart for Weibullpercentiles.QualityReliabilityEngineering Int., 22: 141-151

Oguntunde, P. E., Balogun, O. S, Okagbue, H. I. and Bishop, S. A. (2015). The WeibullExponential Distribution: Its Properties and Applications. Journal of Applied Sciences, 15(11), 1305-1311.

Smith, R.L. and Naylor, (1987). A Comparison of Maximum Likelihood and Bayesian Estimators for the Three-Parameter Weibull Distribution. Journal of Applied Statistics, 36, 358-369.(1987)

Terna G. I (2017). "AWeibull-Normal Distribution: Its Properties and Applications. Unpublish thesis Department of Statistics, Ahmadu Bello University Kaduna.p57

Weibull, W. (1979). A statistical study of the strength of material. VetenskapsAkad. Handl. (Stockholm) Vol. 151, pp. 15 (as cited by Haan, 1986).